American Journal of Medical Genetics 64:209–215 (1996)

Fragile X Founder Chromosomes in Italy: A Few Initial Events and Possible Explanation for Their Heterogeneity

P. Chiurazzi, M. Genuardi, L. Kozak, M.L. Giovannucci-Uzielli, C. Bussani, F. Dagna-Bricarelli, M. Grasso, L. Perroni, G. Sebastio, M.P. Sperandeo, B.A. Oostra, and G. Neri

Istituto di Genetica Medica, Università Cattolica and Centro Ricerche per la Disabilità Mentale e Motoria, Associazione Anni Verdi, Roma (P.C., M.G., L.K., G.N.); Servizio di Genetica Umana, Dipartimento di Pediatria, Università di Firenze, Firenze (M.L.G.-U., C.B.); Centro di Genetica Umana, E.O. Ospedali Galliera, Genova (F.D.-B., M.G., L.P.); Dipartimento di Pediatria, Università Federico II, Napoli (G.S., M.P.S.), Italy; Department of Clinical Genetics, Erasmus University, Rotterdam, The Netherlands (P.C., B.A.O.)

A total of 137 fragile X and 235 control chromosomes from various regions of Italy were haplotyped by analyzing two neighbouring marker microsatellites, FRAXAC1 and DXS548. The number of CGG repeats at the 5′ end of the FMR1 gene was also assessed in 141 control chromosomes and correlated with their haplotypes. Significant linkage disequilibrium between some "major" haplotypes and fragile X was observed, while other "minor" haplotypes may have originated by subsequent mutation at the marker microsatellite loci and/or recombination between them. Recent evidence suggests that the initial mechanism leading to CGG instability might consist of rare ($10^{-6/-7}$) CGG repeat slippage events and/or loss of a stabilizing AGG via A-to-C transversion. Also, the apparently high variety of fragile X chromosomes may be partly due to the relatively high mutation rate ($10^{-4/-5}$) of the microsatellite markers used in haplotyping. Our fragile X sample also showed a higher than expected heterozygosity when compared to the control sample and we suggest that this might be explained by the chance occurrence of the few founding events on different chromosomes, irrespective of their actual frequency in the population. Alternatively, a local mechanism could enhance the microsatellite mutation rate only on fragile X chromosomes, or fragile X mutations might occur more frequently on certain background haplotypes. © 1996 Wiley-Liss, Inc.

KEY WORDS: fragile X, population genetics, founder chromosomes, microsatellite markers, marker heterozygosity

INTRODUCTION

Short tandem repeats (STRs), or microsatellites, are polymorphic head-to-tail arrangements of DNA sequences up to about 6 bp [Weber and Wong, 1993]. They are extremely diffuse in the human genome and are almost evenly distributed between coding and uncoding regions [Tautz and Renz, 1984], thus becoming useful in establishing high density linkage maps [Weissenbach et al., 1992; Gyapay et al., 1994]. They appear to be 5–10 times more frequent than would be expected by chance [Tautz et al., 1986], possibly as a consequence of a slipped-strand mispairing mechanism [Levinson and Gutman, 1987] particularly active in eukaryotic genomes [Charlesworth et al., 1994]. This mechanism has been shown to operate both in vitro [Schloetterer and Tautz, 1992; Hauge and Litt, 1993] and in vivo with mean mutation rates estimated at 10^{-4} to almost 10^{-2} per locus per gamete per generation, although the frequency of slippage seems to be higher in the male compared to the female germline [Weissenbach et al., 1992; Weber and Wong, 1993]. Mutation rates of tetranucleotide repeats 4 times higher than that of dinucleotide repeats has been reported [Mahtani and Willard, 1993; Weber and Wong, 1993]. Trinucleotide repeats also can be tolerated inside transcribed and translated sequences without affecting the reading frame, but their potential instability can result in gene malfunction if excessive amplification occurs [Kuhl and Caskey, 1993]. This is the case of the fragile X syndrome and several neurodegenerative disorders [Willems, 1994]. Although "mutator" genes, such

Received for publication October 20, 1995.

Address reprint requests to Giovanni Neri, MD, Istituto di Genetica Medica, Università Cattolica, L.go F. Vito 1, 00168 Rome, Italy.

L. Kozak's present address is Institute for Child Health, Brno, Czech Republic.

as MSH2, have been identified that can affect the overall genomic slippage rate by altering the efficiency of the mismatch repair systems [Kunkel, 1993], all presently identified "dynamic mutations" seem to be local phenomena, apparently determined by the particular structure of their trinucleotide repeat. Considering the high mutation rate of the CGG repeat at the 5′ untranslated region of the FMR1 gene, whose amplification is responsible for most cases of fragile X syndrome, the finding of linkage disequilibrium between the mutation and certain haplotypes defined by flanking microsatellite markers was unpredicted [Richards et al., 1992; Oudet et al., 1993a,b; Buyle et al., 1993; Hirst et al., 1993; Jacobs et al., 1993; Macpherson et al., 1994; Malmgren et al., 1994; Montagnon et al., 1994; Zhong et al., 1994a]. A founder effect suggests a lengthy multistep process leading to the pathogenetic amplification with different transition rates between each mutational stage. In order to search for the existence of founder chromosomes in Italy, 137 unrelated fragile X and 235 control chromosomes were haplotyped with two neighbouring microsatellite markers, DXS548 and FRAXAC1, and the CGG repeat number was assessed in 141 of the control chromosomes to check for the possible existence of a correlation between total repeat length and "at risk" haplotypes.

PATIENTS AND METHODS

DNA aliquots of 137 unrelated fragile X male patients were collected at four genetic centers (Rome, Genua, Naples, and Florence) thus providing a representative sample of most Italian regions. DNA from 141 healthy males and 47 healthy females was also collected from the same centers and constituted our pool of 235 control chromosomes. The microsatellites employed for haplotyping were FRAXAC1 [Richards et al., 1991] and DXS548 [Riggins et al., 1992] which are located 7 kbp and 150 kbp, respectively, proximal to the CGG repeat at the 5′ end of the FMR1 gene. A duplex PCR protocol was developed to allow simultaneous amplification in a single reaction. Thirty cycles (94°/1 min-55°/1 min-72°/1 min) were performed in a hot bonnet Peltier effect thermal cycler and reactions were carried out in 10 μl with 200 ng genomic DNA, 0.3 units of Taq polymerase, 1.5 mM $MgCl_2$, 10% DMSO, 0.2 mM each dNTP, except dCTP, which was 2 μM, 0.1 μl of alpha-^{32}P-dCTP (3,000 Ci/ml), 10 pmol of each FRAXAC1 [Zhong et al., 1993] and DXS548 [Verkerk et al., 1991] primer. Primer sequences were, respectively, FRAXAC1-F (GAT CTA ATC AAC ATC TAT AGA CTT TAT T), FRAXAC1-R (GAT GAG AGT CAC TTG AAG CTG G), DXS548-F (GTA CAT TAG AGT CAC CTG TGG TGC) and DXS548-R (AGA GCT TCA CTA TGC AAT GGA ATC). A new forward DXS548 primer (GAA TAG TCT CTG GGG TGG ATC TC) has been designed which produces a shorter PCR product including only the main GT repeat (see Results). The DNA of the 141 control males was also amplified with primers "c" and "f" [Fu et al., 1991] to assess the CGG repeat number with 200 ng genomic DNA, 0.65 units of Taq polymerase, 2.0 mM $MgCl_2$, 10% DMSO, 0.2 mM each dNTP, except dGTP which was 0.05 mM, 0.15 mM 7-deaza-dGTP, 0.13 μl of alpha-^{32}P-dCTP (3000 Ci/ml) and 2 pmol of each primer. Thirty cycles were again used in a two-step PCR (94°/45 sec-68°/2 min 30 sec). A 10 μl aliquot of stop solution (95% formamide and dyes) was added to each tube, PCR products were denatured for 10 min at 95° and 4 μl were loaded on a 6% polyacrylamide/7M urea denaturing gel. Samples were run at 70 W for 2–4 h depending on the size of the amplified products (approximately 200 and 150 bp for the duplex PCR and 300 bp for the 30 CGG repeat allele). Gels were dried without fixation and exposed for 12 to 48 hr at −80°C with intensifying screens. Reference DNAs were consistently re-loaded at the sides and in the middle of every gel to rule out gel distortion artifacts and samples were always run a second time in increasing size in order to cross-check their relative size. Some samples were reamplified if they failed to give any signal the first time or if they showed ambiguous bands. CGG repeat size was evaluated next to an M13 sequence assuming the thickest band to be the original product, as 2–3 faster shadow bands were always present [Hauge and Litt, 1993]. FRAXAC1 alleles were named with letters (A–F) as in Richards et al. [1992] and DXS548 alleles were named as in Macpherson et al. [1994], although the exact size of allele 8 calculated by sequencing (187 bp with 16 GT repeats) did not correspond to that (192 bp) quoted by Riggins et al. [1992] and Oudet et al. [1993a]. The chi-square test with Yate's correction is well suited to analyze the results and has been employed to detect significant differences between the fragile X and control samples. Heterozygosity and its variance were estimated using formulas [8.4] and [8.12] described in Nei [1987].

RESULTS

FRAXAC1 and DXS548 allele distributions in both fragile X subjects and controls are shown in Figure 1a and b, respectively. As in all previous studies on Caucasian populations, the most frequent alleles in the controls are FRAXAC1-C [Richards et al., 1992; Hirst et al., 1993; Jacobs et al., 1993; Macpherson et al., 1994; Zhong et al., 1994a] and DXS548-7 [Oudet et al., 1993a,b; Buyle et al., 1993; Malmgren et al., 1994; Montagnon et al., 1994; Macpherson et al., 1994; Zhong et al., 1994a]. The frequency of both alleles decreases significantly in the fragile X population (chi-square 29.8 and 68.5, respectively, $P < 0.001$), matched by an increase in the frequency of other alleles, especially FRAXAC1-A (chi-square 41.2, $P < 0.001$) and DXS548-2 (chi-square 37.4, $P < 0.001$). These values demonstrate the existence of a significant linkage disequilibrium between the fragile X mutation and some alleles of the neighbouring microsatellites. DXS548 intermediate allele 6.5, observed only once, positions itself between alleles 6 and 7. PCR has been repeated several times and the product was also run with an automatic sequencer to confirm the intermediate size of this allele. Examining the DXS548 sequence [S.T. Warren, unpublished] approximately 50 bp proximal to the GT repeat, we could notice a $(C)_4G(C)_{11}$ sequence whose longer poly-C tract might be slightly polymorphic

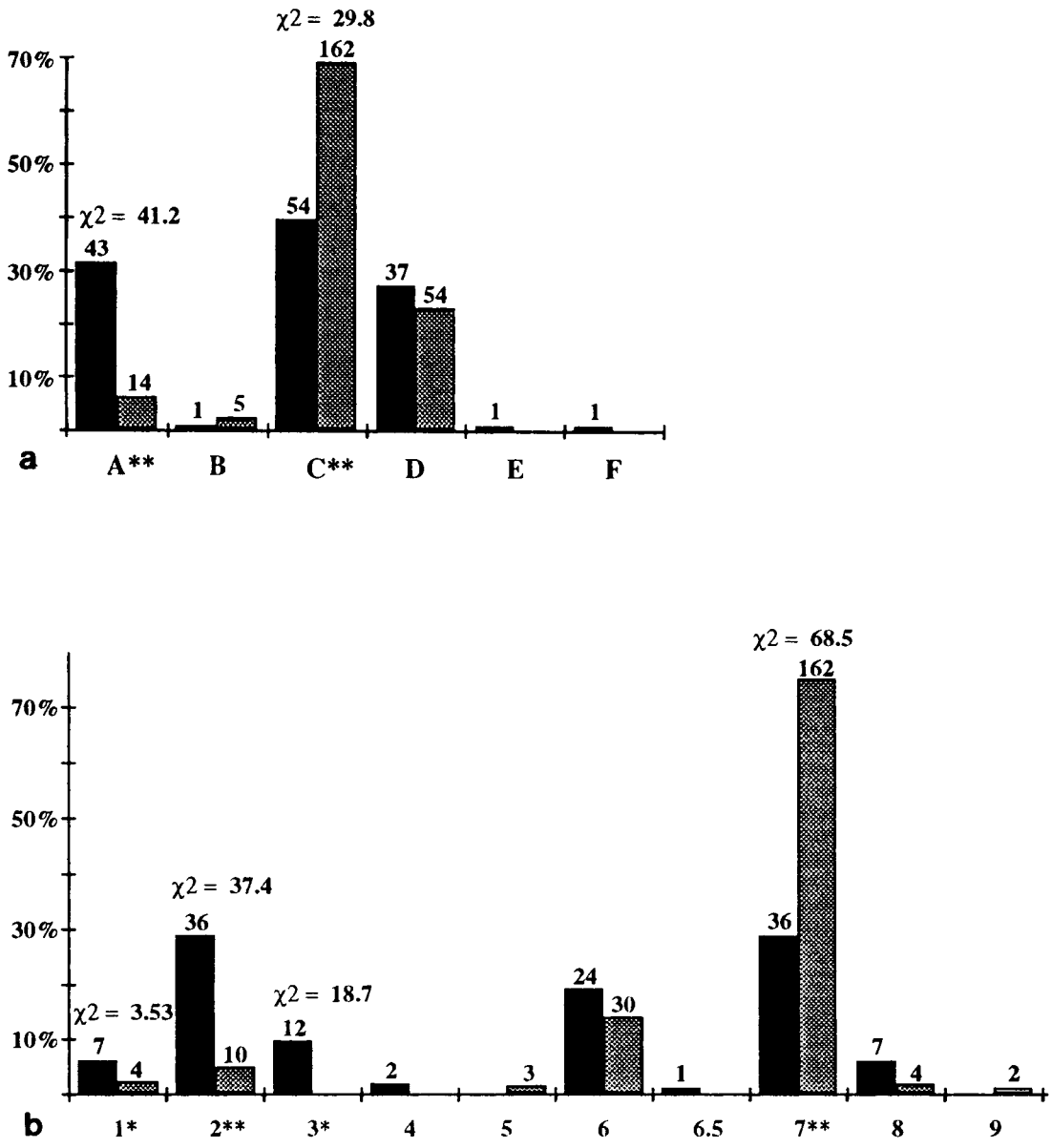

Fig. 1. **a:** Allele distribution at FRAXAC1 locus of 137 fragile X chromosomes (black bars) and 235 controls (gray bars); **b:** allele distribution at DXS548 locus of 125 fragile X chromosomes and 215 controls. Absolute chromosome numbers are reported on the top of every bar. Statistically significant differences are indicated by asterisks next to the allele symbol and chi-square values are written on the top of corresponding bars.

[Weber, 1990]. In fact, when we directly sequenced three clones of a DXS548-8 allele we found a $(C)_4G(C)_{10}$ motif; this 1 bp insertion/deletion polymorphism could account for these intermediate alleles, although its rare occurrence in European populations does not blur the information obtained with the original primers. Anyhow, more of these intermediate alleles have been observed in a normal black African population [Chiurazzi et al., 1996], therefore a new forward DXS548 primer has been designed which gives a shorter (120 bp) product including only the main GT repeat.

Figure 2 represents the FRAXAC1-DXS548 haplotypes in decreasing order of frequency in the control population, and the corresponding frequencies found in the fragile X sample. Again as in previous studies, haplotype C-7 is by far the most common among the controls, while in the fragile X sample its significant decrease (chi-square 55.7, $P < 0.001$) is mainly compensated for by the rise of haplotype A-2 (chi-square 35.2, $P < 0.001$). Also clearly increased in the patients' population are haplotypes C-3 (chi-square 15.8, $P < 0.001$) and D-6 (chi-square 6, $P < 0.015$, while few other haplotypes give a minor contribution to the fragile X pool. A comparison with the data of Macpherson et al. [1994] and, relatively to DXS548 only, with those of Oudet et al. [1993a,b], Buyle et al. [1993] and Malm-

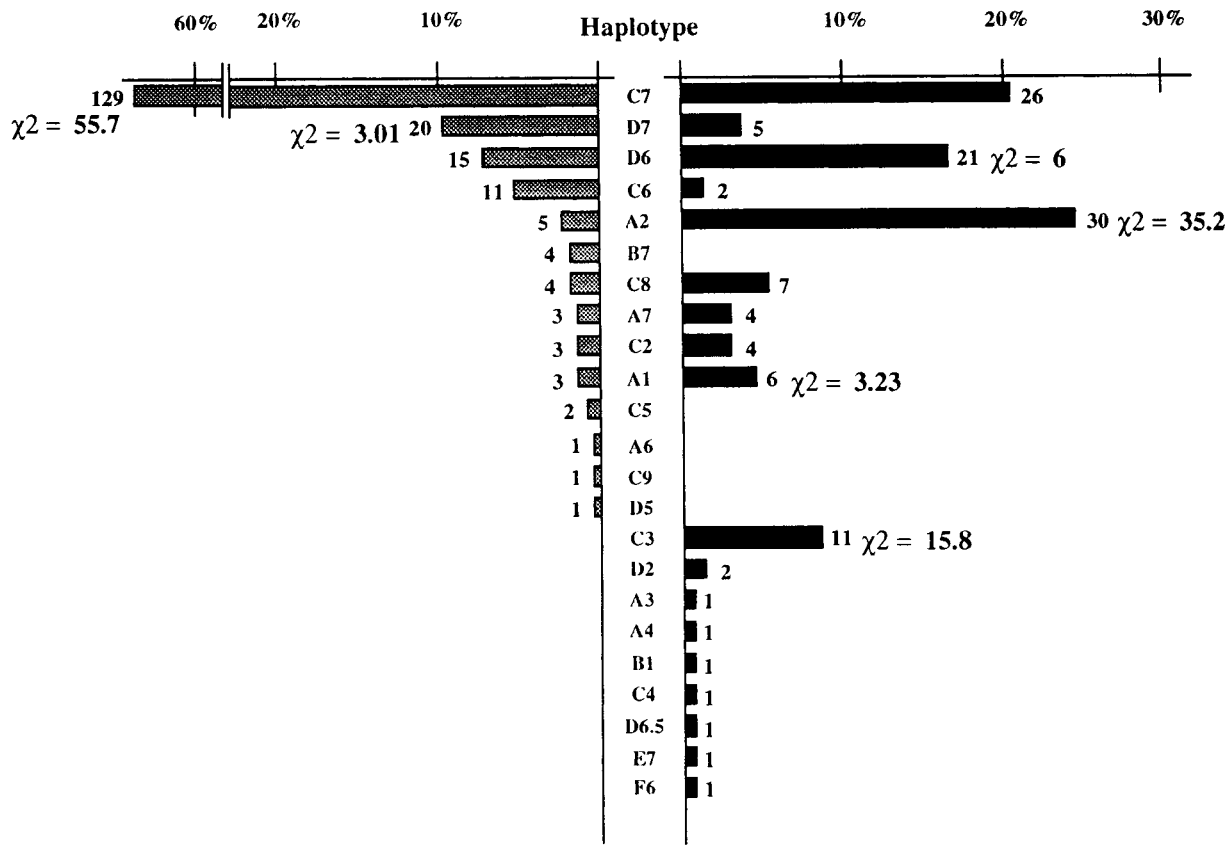

Fig. 2. Haplotype distribution of 125 fragile X chromosomes (black bars) and 202 controls (gray bars) presented in decreasing order of frequency in the normal control population. Absolute chromosome numbers are reported next to every bar. Statistically significant differences are indicated by writing chi-square values beside the corresponding bars.

gren et al. [1994] suggests a striking similarity of fragile X haplotypes distribution, with a South-to-North gradient noted for haplotypes A-2 and D-6, the latter being more prevalent in Scandinavia and Britain, while the former is dominant in Italy and Southern Europe. Actually, A-2 was also observed in 2 of 3 patients from Iran and in one Singhalese boy referred to us [Chiurazzi, unpublished observation]. The CGG repeat distribution of our control population, subdivided according to haplotype, is shown in Figure 3 in order to verify if the total repeat length is a major factor in determining a higher instability of some haplotypes. The few control chromosomes with haplotype A-2 have indeed a higher mean CGG repeat number, as also found by Kunst and Warren [1994]. However, this is not the case for D-6 and the other haplotypes that are more prevalent among the fragile X population, underlining the role of factors other than total repeat length in causing triplet instability.

Figure 4 presents the fragile X and control haplotype distribution as a two-dimensional array of the DXS548 and FRAXAC1 alleles. Its simple inspection allows two important observations: 1) the relatively large number of different fragile X haplotypes, although 4 of them (A-2, C-7, D-6 and C-3), that we shall designate major haplotypes, make up for 70% of the total; 2) the higher heterozygosity (86% ± 1.6%) of the fragile X sample compared to that of controls (57.4% ± 3.9%), which is not unexpected if we consider that the fragile X chromosomes are more evenly distributed over the 4 major haplotypes (A-2, 24%; C-7, 20.8%; D-6, 16.8%; and C-3, 8.8%), while the control distribution has a single dominant peak (C-7, 64%).

DISCUSSION

The study of fragile X population genetics has great relevance to test the various hypotheses made on the nature of the "dynamic" mutational mechanism affecting the FMR1 gene. In fact, the first evidence of linkage disequilibrium between FMR1 mutations and flanking marker loci prompted Morton and Macpherson [1992] to propose a multistep model with the frequency of transition increasing from one mutational stage to the next (N-to-S < S-to-Z < Z-to-L). This basic property is common to other models [Kolehmainen, 1994; Ashley and Sherman, 1995] and allows a few protomutations to generate a relatively high number of full mutations [Chakravarti, 1992]. Thus, founder chromosomes can be identified, although the strongly reduced fitness of the affected males would normally imply many new

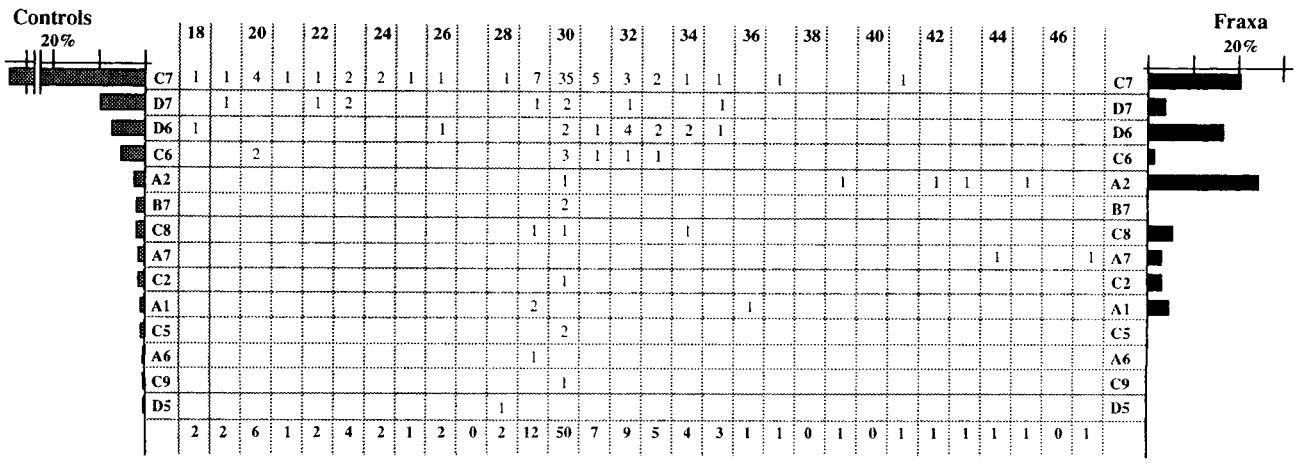

Fig. 3. CGG repeat number distribution of 141 male controls, subdivided according to haplotype.

mutations on almost every chromosomal background [Sherman et al., 1985] if a single-step mutational mechanism is assumed.

A complex mechanism also emerges from recent observations on slipped-strand mispairing and hairpin formation, delineating a double-stage scenario. A first phase of low-grade instability would include the eventual loss of stabilizing imperfections as intercalated AGGs [Kunst and Warren, 1994; Eichler et al., 1994; Hirst et al., 1994; Snow et al., 1994; Zhong et al., 1995] and slippage of a few or, more rarely, even 10 repeat units [Macpherson et al., 1995]. Such a complex stepwise mutational mechanism (SMM), possibly common to all microsatellites, would allow both expansions and reductions to occur even in discrete "jumps" and seems to fit better the experimental data [Di Rienzo et al., 1994; Deka et al., 1995; Watkins et al., 1995]. The determinant factors for this initial instability would be

DXS548

FRAXACI	1	2	3	4	5	6	6.5	7	8	9	
A	4.80* 1.49	24.00** 2.48	0.80	0.80		0.50		3.20 1.49			33.60 5.94
B	0.80							1.98			0.80 1.98
C		3.20 1.49	8.80**	0.80	0.99	1.60 5.45		20.80* 63.86*	5.60 1.98	0.50	40.80 74.26
D		1.60			0.50	16.80* 7.43	0.80	4.00 9.90*			23.20 17.82
E								0.80			0.80
F						0.80					0.80
	5.60 1.49	28.80 3.96	9.60	1.60	1.49	19.20 13.37	0.80	28.80 77.23	5.60 1.98	0.50	100 100

Fig. 4. Haplotype distribution of 125 fragile X chromosomes (bold type) and 202 controls presented as a bidimensional array. Frequencies rather than numbers are given in order to facilitate comparison and statistically significant differences are indicated by asterisks. Haplotype expected heterozygosity (± standard error) was estimated for the fragile X (86% ± 1.6%) and for the control samples (57.4% ± 3.9%).

both the CGG repeat internal structure and its length, as also suggested by our data presented in Figure 3. Then, a second phase of more pronounced instability, leading to larger expansions, would be possible only for microsatellite with specific repeat sequences because, above a certain length threshold, formation of stable hairpin structures is favoured [Gacy et al., 1995] and, in the case of CGG repeats, hairpins would facilitate and be further stabilized by cytosine methylation [Fry and Loeb, 1994]. Moreover, homogenous Okazaki fragments composed almost exclusively of repeats would be even more prone to mispairing and hairpin formation [Richards and Sutherland, 1994]; while the location and distance of the closest origin of replication might favour expansion over deletion events [Kang et al., 1995] and account for the amplification polarity [Kunst and Warren, 1994]. Kolehmainen [1994] correctly pointed out that an estimate of 2×10^{-4} for the initial transition rate (wild-type to protomutation) would result in too many founder alleles to accept the suggestion that all fragile X mutations originated on a small number of founder chromosomes. We suggest that the above-mentioned mechanistic evidence is compatible with a downward revision of the estimate of the initial (N-to-S) transition rate to values comparable to those of most point mutations ($10^{-6/-7}$), as in the case of the A-to-C transversion leading to the loss of a stabilizing AGG [Eichler et al., 1994]. Furthermore, the striking similarity among Caucasian fragile X haplotypes, the South-North gradient of haplotype A-2 possibly originated on an ancient Indoeuropean chromosome, and the analysis of fragile X pedigrees extending over several generations [Smits et al., 1993], all argue in favour of an extremely long history of fragile X mutations. Indeed, recent studies on a black African population [Chiurazzi et al., 1996] and on primates [Deelen et al., 1994; Rubinsztein et al., 1995; Zhong et al., 1995] show that the CGG repeat at the 5′ end of the FMR1 gene is evolutionarily conserved and already attained the mean size of 30 repeats with stabilizing AGGs in the primates. Of course, some founder chromosomes, as D-6, could be of more recent origin and may have not yet contributed fully to the pool of fragile

X patients yet [Kunst and Warren, 1994; Mandel, 1994]. Based on the above evidence, we may now assume a multistep mutational mechanism with initial transition rate of $10^{-6/-7}$ and reinterpretate the first of the observations derived from Figure 4. The relative large number of different fragile X haplotypes can be reconciled with a strong linkage disequilibrium; in fact only the major haplotypes (A-2, C-7, D-6, C-3, and possibly A-1 and C-8) probably reflect a mutational event and represent the original founder chromosomes, while the other fragile X haplotypes may have originated from the major ones via slippage at one or both flanking markers, given the relatively higher ($10^{-3/-5}$) mutation rate of microsatellites [Weber and Wong, 1993]. In less frequent cases, recombination between microsatellite markers might have created a minor haplotype, one such case possibly being haplotype A-7, which could originate from a crossing-over between a founder A-2 and a common control C-7 haplotype. In fact, the only two A-7 control chromosomes have a high CGG repeat number, as it is the case for most A-2 chromosomes from which they would be derived. The identification and use of less polymorphic markers (e.g., RFLPs) in the FMR1 gene will be crucial for the future assessment of this problem.

The second observation we made from Figure 4 was that of the higher heterozygosity (86% ± 1.6%) of the fragile X sample compared to that of controls (57.4% ± 3.9%). This is not an obvious finding because in the absence of a founder effect, when new mutations are frequent, the haplotype distributions of patients and controls would be almost identical, while in presence of a strong founder effect the majority of mutations would be expected to occur on a more limited haplotype background. However, the observed high heterozygosity could still be compatible with few founding events if they occurred on different and distant haplotypes, irrespective of their relative frequency in the control population. Of course, it would have been more likely for the founding chromosomes to cluster on and around the common control C-7 haplotype, but fragile X mutations during history might have generated more dispersed clusters of major and minor founder chromosomes by chance. This "casual" hypothesis has the advantage of not implying any particular mechanism linking the stability of the FMR1 CGG repeat with that of its neighbouring marker loci, but in alternative, as suggested by Morris et al. [1995], we have to consider three other possible models accounting for the association of some haplotypes with the fragile X mutations. In the "concurrent" model an external factor, perhaps a mutation at a DNA repair locus, causes a simultaneous instability of the fragile X locus and of its flanking markers. Although there has been no observation of increased microsatellite instability at other genomic loci in most fragile X families, this could be the case of a minority of families, possibly including that with multiple FRAXAC2 mutations reported by Zhong et al. [1993]. The "causal" model supposes that some haplotypes might predispose to mutation the FMR1 CGG repeat, but this would imply the presence of yet another cis-element acting on the triplet while it seems that length and purity of the CGG repeat itself are the main determinants of its stability. Finally, the "consequent" model suggests that the fragile X mutation might enhance the neighbouring microsatellite mutation rate [Zhong et al., 1994a,c; Mornet et al., 1994; Richards et al., 1994b; Zhong et al., 1995], and this deserves careful verification although it would tend to reduce any linkage disequilibrium by definition and there is no clear evidence of a plausible mechanism. However, a thorough analysis of larger sets of data from Asiatic populations might prove helpful, considering that Japanese [Richards et al., 1994a] and possibly Chinese populations [Zhong et al., 1994b] are less heterogenous than European ones and present very similar fragile X and control haplotype distributions; this should allow an easier verification of the possible existence of an increased mutability of flanking markers on fragile X chromosomes.

ACKNOWLEDGMENTS

This work was supported in part by contribution from MURST. P.C. wishes to express his gratitude to Dr. Esther de Graaff and Prof. Hans Galjaard for their generous support and gratefully acknowledges the financial support of a C.N.R. grant for his travel and stay in Rotterdam. We are indebted to all the families of our patients for their continuous encouragement and cooperation.

REFERENCES

Ashley AE, Sherman SL (1995): Population dynamics of a meiotic/mitotic expansion model for the Fragile X syndrome. Am J Hum Genet 57:1414–1425.

Buyle S, Reyniers E, Vits L, De Boulle K, Handig I, Wuyts FLE, Deelen W, Halley DJJ, Oostra BA, Willems PJ (1993): Founder effect in a Belgian-Dutch fragile X population. Hum Genet 92: 269–272.

Chakravarti A (1992): Fragile X founder effect? Nat Genet 1:237–238.

Charlesworth B, Sniegowski P, Stephan W (1994): The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220.

Chiurazzi P, Destro-Bisol G, Genuardi M, Oostra BA, Spedini G, Neri G (1996): Extended gene diversity at the FMR1 locus and neighbouring CA repeats in a sub-Saharan population. Am J Med Genet, 64:216–219.

Deelen W, Bakker C, Halley DJJ, Oostra BA (1994): Conservation of CGG region in FMR1 gene in mammals. Am J Med Genet 51: 513–516.

Deka R, Jin L, Shriver MD, Yu LM, DeCroo S, Hundrieser J, Bunker CH, Ferrell RE, Chakraborty R (1995): Population genetics of dinucleotide (CA)(GT) polymorphism in world populations. Am J Hum Genet 56:461–474.

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994): Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci 91:3166–3170.

Eichler EE, Holden JJA, Popovich BW, Reiss AL, Snow K, Thibodeau SN, Richards CS, Ward PA, Nelson DL (1994): Length of uninterrupted CGG repeats determines instability in the FMR1 gene. Nat Genet 8:88–94.

Fry M, Loeb LA (1994): The fragile X syndrome d(CGG)$_n$ nucleotide repeats form a stable tetrahelical structure. Proc Natl Acad Sci USA 91:4950–4954.

Fu YH, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJMH, Holden JJA, Fenwick RG, Warren ST, Oostra BA, Nelson DL, Caskey CT (1991): Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67:1047–1058.

Gacy AM, Goellner G, Juranic N, Macura S, McMurray CT (1995): Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell 81:533–540.

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994): The 1993–1994 Genethon human genetic linkage map. Nat Genet 7: 246–249.

Hauge XY, Litt M (1993): A study of the origin of "shadow bands" seen when typing dinucleotide repeat polymorphisms by the PCR. Hum Mol Genet 2:411–415.

Hirst MC, Knight SJL, Christodoulou Z, Grewal PK, Fryns JP, Davies KE (1993): Origins of the fragile X syndrome mutation. J Med Genet 30:647–650.

Hirst MC, Grewal PK, Davies KE (1994): Precursor arrays for the triplet repeat expansion at the fragile X locus. Hum Mol Genet 3:1553–1560.

Jacobs PA, Bullman H, Macpherson J, Youings S, Rooney V, Watson A, Dennis NR (1993): Population studies of the fragile X: a molecular approach. J Med Genet 30:454–459.

Kang S, Jaworski A, Ohshima K, Wells RD (1995): Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E. coli. Nat Genet 10:213–218.

Kolehmainen K (1994): Population genetics of fragile X: a multiple allele model with variable risk of CGG repeat expansion. Am J Med Genet 51:428–435.

Kuhl DPA, Caskey CT (1993): Trinucleotide repeats and genome variation. Curr Opin Genet Develop 3:404–407.

Kunkel TA (1993): Slippery DNA and diseases. Nature 365:207–208.

Kunst CB, Warren ST (1994): Cryptic polar variation of the fragile X repeat could result in predisposing normal alleles. Cell 77:853–861.

Levinson G, Gutman GA (1987): Slipped strand mispairing: A major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221.

Macpherson JN, Bullman H, Youings SA, Jacobs PA (1994): Insert size and flanking haplotype in fragile X and normal populations: Possible multiple origins for the fragile X mutation. Hum Mol Genet 3:399–405.

Macpherson JN, Curtis G, Crolla JA, Dennis N, Migeon B, Grewal PK, Hirst MC, Davies KE, Jacobs PA (1995): Unusual $(CGG)_n$ expansion and recombination in a family with fragile X and DiGeorge syndrome. J Med Genet 32:236–239.

Mahtani MM, Willard HF (1993): A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. Hum Mol Genet 2:431–437.

Malmgren H, Gustavson KH, Oudet C, Holmgren G, Pettersson U, Dahl N (1994): Strong founder effect for the fragile X syndrome in Sweden. Eur J Hum Genet 2:103–109.

Mandel JL (1994): Trinucleotide diseases on the rise. Nat Genet 7:453–455.

Montagnon M, Boygo A, Deluchat C, Jokic M, Chateau C, Taillandier A, Thomas F, Simon-Bouy B, Boué J, Serre JL, Boué A, Mornet E (1994): Transition from normal to premutated alleles in fragile X syndrome results from a multistep process. Eur J Hum Genet 2:125–131.

Mornet E, Chateau C, Taillandier A, Montagnon M, Simon-Bouy B, Serre JL, Boué A (1994): FRAXAC2 instability. (letter) Nat Genet 7:122–123.

Morris A, Morton NE, Collins A, Lawrence S, Macpherson JN (1995): Evolutionary dynamics of the FMR1 locus. Ann Hum Genet 59:283–289.

Morton NE, Macpherson JN (1992): Population genetics of the fragile X syndrome: Multiallelic model for the FMR1 locus. Proc Natl Acad Sci USA 89:4215–4217.

Nei M (1987): Molecular evolutionary genetics. New York: Columbia University Press.

Oudet C, Mornet E, Serre JL, Thomas F, Lentes-Zengerling S, Kretz C, Deluchat C, Tejada I, Boué J, Boué A, Mandel JL (1993a): Linkage disequilibrium between the fragile X mutation and two closely linked CA repeats suggest that fragile X chromosomes are derived from a small number of founder chromosomes. Am J Hum Genet 52:297–304.

Oudet C, von Kiskull H, Nordstroem AM, Peippo M, Mandel JL (1993b): Striking founder effect for the fragile X syndrome in Finland. Eur J Hum Genet 1:181–189.

Richards RI, Sutherland GR (1994): Simple repeat DNA is not replicated simply. Nat Genet 6:114–116.

Richards RI, Holman K, Kozman H, Kremer E, Lynch M, Pritchard M, Yu S, Mulley J, Sutherland GR (1991): Fragile X syndrome: Genetic localization by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. J Med Genet 28:818–823.

Richards RI, Holman K, Friendendend K, Kremer E, Hillen D, Staples A, Brown WT, Goonewardena P, Tarleton J, Schwartz C, Sutherland GR (1992): Evidence of founder chromosomes in fragile X syndrome. Nat Genet 1:257–260.

Richards RI, Kondo I, Holman K, Yamauchi M, Seki N, Kishi K, Staples A, Sutherland GR, Hori T (1994a): Haplotype analysis at the FRAXA locus in the Japanese population. Am J Med Genet 51:412–416.

Richards RI, Holman K, Friend K, Staples A, Sutherland GR, Oudet C, Biancalana V, Mandel JL (1994b): FRAXAC2 instability. (letter) Nat Genet 7:122.

Riggins GJ, Sherman SL, Oostra BA, Sutcliffe JS, Feitell D, Nelson DL, van Oost BA, Smits APT, Tamos FJ, Pfender E, Kuhl DPA, Caskey CT, Warren ST (1992): Characterization of a highly polymorphic dinucleotide repeat 150 kb proximal to the fragile X site. Am J Med Genet 43:237–243.

Rubinsztein DC, Amos W, Leggo J, Goodburn S, Jain S, Li SH, Margolis RL, Ross CA, Ferguson-Smith MA (1995): Microsatellite evolution-evidence for directionality and variation in rate between species. Nat Genet 10:337–343.

Schloetterer C, Tautz D (1992): Slippage synthesis of simple sequence DNA. Nucl Ac Res 20:211–215.

Sherman SL, Jacobs PA, Morton NE, Froster-Iskenius U, Howard-Peebles PN, Nielsen KB, Partington MW, Sutherland GR, Turner G, Watson M (1985): Further segregation analysis of the fragile X syndrome with special reference to transmitting males. Hum Genet 69:289–299.

Smits APT, Dreesen JCFM, Post JG, Smeets DFCM, de Die-Smulders C, Spaans-van der BiJl T, Govaerts LCP, Warren ST, Oostra BA, van Oost BA (1993): The fragile X syndrome: no evidence for any recent mutations. J Med Genet 30:94–96.

Snow K, Tester DJ, Kruckeberg KE, Schaid DJ, Thibodeau SN (1994): Sequence analysis of the fragile X trinucleotide repeat: implications for the origin of the fragile X mutation. Hum Mol Genet 3: 1543–1551.

Tautz D, Renz M (1984): Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucl Ac Res 12:4127–4138.

Tautz D, Trick M, Dover GA (1986): Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652–656.

Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen GJB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL, Oostra BA, Warren ST (1991): Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65:905–914.

Watkins WS, Bamshad M, Jorde LB (1995): Population genetics of trinucleotide repeat polymorphism. Hum Mol Genet 4:1485–1491.

Weber JL (1990): Informativeness of human $(dC\text{-}dA)_n\cdot(dG\text{-}dT)_n$ polymorphism. Genomics 7:524–530.

Weber JL, Wong C (1993): Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128.

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992): A second-generation linkage map of the human genome. Nature 359:794–801.

Willems PJ (1994): Dynamic mutations hit double figures. Nat Genet 8:213–215.

Zhong N, Dobkin C, Brown WT (1993): A complex mutable polymorphism located within the fragile X gene. Nat Genet 5:248–252.

Zhong N, Ye L, Dobkin C, Brown WT (1994a): Fragile X founder chromosome effects: Linkage disequilibrium or microsatellite heterogeneity? Am J Med Genet 51:405–411.

Zhong N, Liu X, Gou S, Houck GE Jr, Li S, Dobkin C, Brown WT (1994b): Distribution of FMR1 and associated microsatellite alleles in a normal Chinese population. Am J Med Genet 51:417–422.

Zhong N, Dobkin C, Brown WT (1994c): FRAXAC2 instability. (letter) Nat Genet 7:123.

Zhong N, Yang W, Dobkin C, Brown WT (1995): Fragile X gene instability: Anchoring AGGs and linked microsatellites. Am J Hum Genet 57:351–361.